# Truthing the stretch: Non-perturbative cosmological realizations with multiscale spherical collapse

Mark C. Neyrinck[1]

[1]*Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA*

14 October 2015

**ABSTRACT**

Here we present a simple, parameter-free, non-perturbative algorithm that gives low-redshift cosmological particle realizations accurate to few-Megaparsec scales, called MUSCLE (MUltiscale Spherical ColLapse Evolution). It has virtually the same cost as producing $N$-body-simulation initial conditions, since it works with the 'stretch' parameter $\psi$, the Lagrangian divergence of the displacement field. It promises to be useful in quickly producing mock catalogs, and to simplify computationally intensive reconstructions of galaxy surveys. MUSCLE applies a spherical-collapse prescription on multiple Gaussian-smoothed scales. It achieves higher accuracy than perturbative schemes (Zel'dovich and 2LPT), and, by including the void-in-cloud process (voids in large-scale collapsing regions), solves problems with a single-scale spherical-collapse scheme. Slight further improvement is possible by mixing in the 2LPT estimate on large scales. Additionally, we show the behavior of $\psi$ for different morphologies (voids, walls, filaments, and haloes). A Python code to produce these realizations is available at http://skysrv.pha.jhu.edu/~neyrinck/muscle.html.

**Key words:** large-scale structure of Universe – cosmology: theory

## 1 INTRODUCTION

In the current paradigm, structures on extragalactic scales arise from the stretching and bunching-together of the 'dark-matter sheet.' A conceptual understanding of this process is a fundamental goal of astronomy, and is closely tied to several other topics in cosmology and galaxy formation. A Lagrangian picture, following particles (i.e. mesh locations on this 'dark-matter sheet') (Shandarin et al. 2012; Abel et al. 2012) instead of fixed comoving positions, is essential to understanding this process. Approximations to full $N$-body dynamics are useful both for this understanding, and to produce approximate $N$-body realizations, e.g. for mock galaxy catalogs (e.g. Kitaura et al. 2014), and for fast explorations of large parameter spaces for Bayesian inference of the initial-conditions (IC) density field (e.g. Kitaura & Enßlin 2008; Kitaura et al. 2012; Kitaura 2013; Heß et al. 2013; Jasche & Wandelt 2013; Leclercq et al. 2015).

The Zel'dovich approximation (Zel'dovich 1970, ZA), linear-order LPT (Lagrangian perturbation theory) already captures much of the essential physics of structure formation (e.g. White 2014), producing a cosmic web accurately to rather small scales. Improvements on ZA in various regimes have been proposed. Going to second order (2LPT) improves accuracy for small fluctuations, i.e. at large scales and early times, useful for IC generation (Scoccimarro 1998; Crocce et al. 2006). At small scales, truncating the power spectrum shortward of the nonlinear scale (Kofman et al. 1992), or using the adhesion model (Kofman & Shandarin 1988; Kofman et al. 1992; Shandarin 2009; Valageas & Bernardeau 2011; Hid-

ding et al. 2015) suppresses unphysical overcrossing of particles in collapsed structures.

Many of these approaches work with the Lagrangian divergence of the displacement field, $\psi(\mathbf{q}) \equiv \boldsymbol{\nabla}_q \cdot \boldsymbol{\Psi}(\mathbf{q})$. Here, the displacement field is $\boldsymbol{\Psi}(\mathbf{q}) = \mathbf{x}(\mathbf{q}) - \mathbf{q}$, where $\mathbf{q}$ denotes Lagrangian (IC) coordinates of a mass element, and $\mathbf{x}$ denotes its Eulerian, final coordinates. A recent approach (Neyrinck 2013, N13) works non-perturbatively with $\psi$, using a low-density limit of the spherical-collapse (SC) evolution of a mass element, found by Bernardeau (1994). This SC relationship was also investigated by Protogeros & Scherrer (1997) and in the context of IC reconstructions by Mohayaee et al. (2006).

Importantly, displacement-divergence ($\psi$-based) schemes giving cosmological realizations are essentially as fast as producing the initial conditions for an $N$-body simulation. They have three steps: (1) Generate a pixelated linear-theory density field $\delta_{\rm lin}$ at the desired redshift, consistent with a linear power spectrum. (2) Estimate $\psi$ from $\delta_{\rm lin}$. (3) Generate the final displacement field $\boldsymbol{\Psi}$ with an inverse-divergence operator, e.g. using an FFT; apply $\boldsymbol{\Psi}$ to particles on a regular lattice to get their final positions.

In the ZA, $\psi_{\rm Z}(\delta_{\rm lin}) = -\delta_{\rm lin}$; in 2LPT, there is a non-local functional giving $\psi$, roughly parabolic in the local $\delta_{\rm lin}$ (N13). In the SC prescription, for each particle (IC density field pixel),

$$\psi_{\rm sc}(\delta_{\rm lin}) = 3\sqrt{1 - (2/3)\delta_{\rm lin}} - 3, \delta_{\rm lin} < 3/2;$$
$$= -3, \ \delta_{\rm lin} \geqslant 3/2. \tag{1}$$

Or, (over)compactly,

$$\psi_{\rm sc}(\delta_{\rm lin}) \equiv 3\,{\rm Re}\sqrt{1 - (2/3)\delta_{\rm lin}} - 3. \tag{2}$$

Setting $\psi_{\mathrm{sc}} = -3$ produces a collapsed volume element as closely as possible when dealing only with $\psi$, since $\psi = -3$ in three dimensions if adjacent particles coincide.

SC successfully reins in particles, preventing overcrossing on the interparticle scale, even when $\delta_{\mathrm{lin}}$ is non-perturbatively large. It also gets densities remarkably right in voids, unlike ZA (overevacuating them), and 2LPT (underevacuating them) (Sahni & Shandarin 1996). When applied far enough outside the perturbative regime of small fluctuations, 2LPT even produces overdensities in void centers (N13). However, when applied at the single, highest Lagrangian resolution of a nonlinear-resolution density field, SC has problems as well. It gives realizations with low power on large scales; also, the Fourier-space cross-correlation coefficient departs from unity at larger scales than in ZA.

Several other techniques have been proposed recently to produce approximate, fast $N$-body or halo density fields, such as PINOCCHIO (Monaco et al. 2002, 2013), COLA (Tassev et al. 2013), and QPM (White et al. 2014). Kitaura & Heß (2013, KH13) found one insightful fix to the SC's issues, calling it Augmented Lagrangian Perturbation Theory (ALPT). In Fourier space,

$$\psi_{\mathrm{ALPT}}(k) = G(k)\psi_{\mathrm{2LPT}}(k) + [1 - G(k)]\psi_{\mathrm{sc}}(k), \qquad (3)$$

an interpolation in scale between 2LPT, which excels on large scales with small fluctuations, and SC, which gets small scales right. Here $G(k)$ is a Gaussian smoothing kernel. ALPT is a key ingredient in the PATCHY algorithm (Kitaura et al. 2014), which produces mock galaxy catalogs using some other novel, useful prescriptions to treat redshift-space distortions and galaxy biasing.

## 2    METHOD AND RESULTS

Here, we give an alternative fix to the SC prescription: making it multiscale. As we show below, the trouble with the SC prescription is that it is applied at the single scale of the interparticle spacing, preventing voids in clouds (voids within larger-scale collapsing regions, e.g. Sheth & van de Weygaert 2004) from properly collapsing.

MUltiscale Spherical-ColLapse Evolution (MUSCLE) additionally checks scales larger than the interparticle scale for collapses, thus including the void-in-cloud process. Like the above schemes, MUSCLE works with step (2), the estimation of $\psi(\delta_{\mathrm{lin}})$, with $\delta_{\mathrm{lin}}$ defined on an interparticle separation (resolution of the initial density field), $R_{\mathrm{ip}}$. The prescription for $\psi$ is

$$\psi_{\mathrm{musc}}(\delta_{\mathrm{lin}}) = 3\left(\sqrt{1 - \frac{2}{3}\delta_{\mathrm{lin}}} - 1\right), \qquad \delta_{\mathrm{lin}} < 3/2 \text{ and}$$
$$G_r(\delta_{\mathrm{lin}}) < 3/2, \forall r \geqslant R_{\mathrm{ip}};$$
$$= -3, \text{ otherwise.} \qquad (4)$$

$G_r(\delta_{\mathrm{lin}})$ denotes the $\delta_{\mathrm{lin}}$ field, Gaussian-filtered at scale $r$. Note that generally, $\psi_{\mathrm{musc}}$ will have nonzero mean, so an additional step is to subtract a constant from all $\psi_{\mathrm{musc}} > -3$ to give zero mean.

In practice, a finite, and preferably small, number of $r > R_{\mathrm{ip}}$ have to be tried. As implemented, we search for collapses at $r = 2^n R_{\mathrm{ip}}$, for integers $n \geqslant 0$. The Gaussian smoothing is done in Fourier space, requiring an FFT and inverse FFT. We also tried using a cubic top-hat filter, averaging together sets of 8 particles on larger scales. This did visually localize halo regions better than the Gaussian smoothing, but the below cross-correlation performed more poorly.

A large factor between smoothing scales searched for collapses gives greater speed, while a small factor captures more collapses, giving more accuracy. At our fiducial factor of 2, generating a random $\Lambda$CDM $256\,h^{-1}$ Mpc realization with $256^3$ particles at $z = 0$ with the MUSCLE Python code searched 7 different smoothing scales; after no collapses were found at $64\,h^{-1}$ Mpc, no larger scales were necessary. On one 2.6 GHz processor, generating this realization with ZA, 2LPT, and MUSCLE took 14, 28, and 59 seconds, respectively.

Fig. 1 shows 2D patches of 3D particle realizations at redshift $z = 0$ using various approximate schemes from the initial conditions. It also shows the full $N$-body results, and 'Perfect $\psi$' – this is generated by measuring the actual $\psi$ field from each particle's $N$-body position using an FFT, and inserting it into step 2 above. This procedure zeroes the curl of the displacement field, but preserves all information in the displacement-divergence $\psi$. See Chan (2014) for another investigation of zeroing the curl. In Fig. 2, a red line is drawn from each particle's actual $N$-body final position to that in the approximate scheme.

In voids, densities (visible from particle separations) are too low in the ZA (Zeld), and too high in 2LPT. Haloes are puffy in Zeld, and even puffier in 2LPT. In this simulation, the interparticle scale is about $0.8\,h^{-1}$ Mpc, with $256^3$ particles in a $200\,h^{-1}$ Mpc box. The linear-theory variance on this scale much exceeds unity, $\sigma_{\mathrm{lin}}^2 = 7.3$; it is not surprising that a higher-order perturbative scheme (2LPT) fails worse than a lower-order scheme (Zeld) when the perturbative parameter is large. SC, as implemented in Eq. (1), gets densities right in voids, and pulls in overcrossed particle positions in haloes, but as Fig. 2 shows, rather large-scale displacements are wrong in SC. In MUSCLE, the large-scale displacement problems are largely solved. This shows that the major deficiency in SC is in its treatment of the void-in-cloud process.

We also show results of 'MUSCLE+2LPT', using the ALPT strategy of interpolating $\psi$ between $\psi_{\mathrm{2LPT}}$ on large scales and $\psi_{\mathrm{MUSCLE}}$ on small scales, as in Eq. 3. The best interpolation smoothing parameter using MUSCLE at $z = 0$ seems to be about $30\,h^{-1}$ Mpc, judging by $R(k)$ as shown below in Fig. 3. We also observed gains in $R(k)$ in ALPT (using SC instead of MUSCLE) on small scales up to this smoothing length, although KH13 found that a smaller smoothing was optimal. So, we use the same smoothing length for each. Interpolating MUSCLE with 2LPT gives a few-percent boost in $R(k)$ on non-linear scales, and gives perhaps a slight visual improvement over MUSCLE, in Figs. 1 and 2. A smaller interpolation scale, $\lesssim 10\,h^{-1}$ Mpc, resulted in 2LPT visibly corrupting large haloes and voids.

In Fig. 3, we show measures of the agreement between the $N$-body result and various approximations. At top, we show the Fourier-space cross correlation, $R(k) = P_{\delta \times \delta'}/\sqrt{P_\delta P_{\delta'}}$ between the $z = 0$ density field and the various approximate density fields, using nearest-gridpoint density assignment on a $128^3$ grid. $P_{\delta \times \delta'}$ here is the cross-power between $\delta$ and $\delta'$. Judging by $R(k)$, sensitive to both Fourier amplitudes and phases, 2LPT performs a couple of per cent better than MUSCLE at $k \sim 0.2\,h^{-1}$ Mpc. Thus, even though MUSCLE's detection of large-scale collapses somehow manages to treat the large-scale tidal field with some accuracy, 2LPT still seems to be the best available method for small fluctuations, e.g. to generate high-redshift initial conditions.

However, MUSCLE outperforms both perturbative schemes on small scales. Interpolating in scale between 2LPT and MUSCLE ('MUSCLE+2LPT') gives the highest $R(k)$ at all scales. Interpolating between 2LPT and SC (labeled 'ALPT') gives results only slightly worse than interpolating with MUSCLE. Note that due to
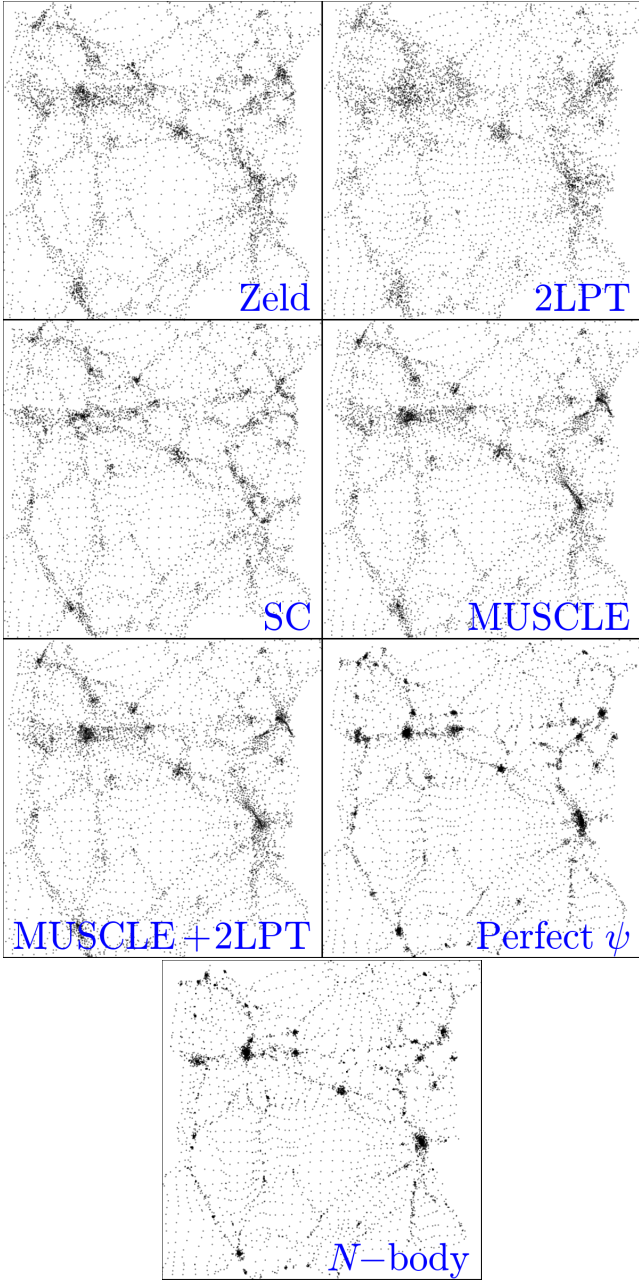
**Figure 1.** Initially square 2D patches, $60\,h^{-1}$ Mpc on a side, of $256^3$-particle realizations made using approximate schemes using the displacement-divergence $\psi$. In 'MUSCLE+2LPT,' $\psi_{\mathrm{MUSCLE}}$ is interpolated in scale as in ALPT(shown here as well, using SC instead of MUSCLE), both with smoothing lengths of $30\,h^{-1}$ Mpc. 'Perfect $\psi$' shows positions after removing the curl of the displacement field, and shows '$N$-body' shows the actual simulation.



**Figure 2.** Same as Fig. 1, except with a red line joining the position of each particle in the full $N$-body result to the approximation.

cosmic variance, small differences in $R(k)$ may not persist for all realizations (see KH13 for error estimates.) However, note that the $R(k)$ curves may (co)vary together up and down more than independently, so small differences may be more statistically significant than the error bars suggest. The superiority of all Lagrangian schemes is obvious from the low 'Eulerian linear theory' curve (Tassev & Zaldarriaga 2012), essentially the propagator (Crocce & Scoccimarro 2006).

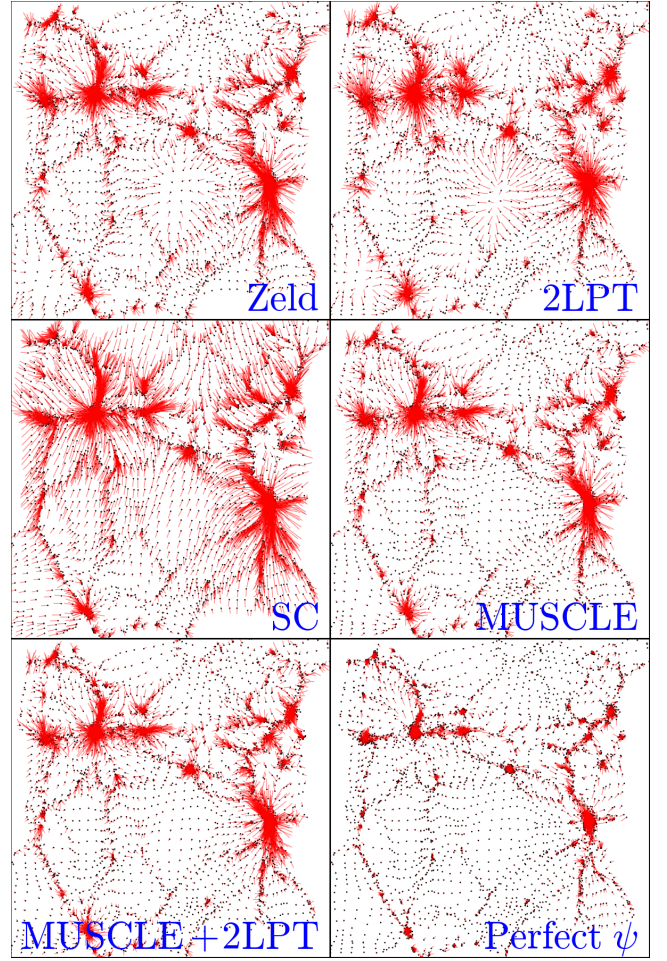The bottom panel of Fig. 3 shows the ratio of the nonlinear power spectrum to those in the various approximations. The power spectra in the $\psi$-based schemes differd only slightly, but the approximation with power closest to full $N$-body is 'MUSCLE+2LPT.' Notably, the large-scale power deficiency in SC is fixed in MUSCLE. In both measurements, while MUSCLE performs well, there is still much distance left to go to 'Perfect $\psi$,' which would be the result if the final $\psi$ could be perfectly predicted.

MUSCLE also gets the 1-point PDF of the field far more accurately than the perturbative approaches. Fig. 4 shows mass-weighted histograms of densities computed using the Voronoi Tessellation Field Estimator (e.g. Schaap & van de Weygaert 2000), with each particle contributing once to the histogram. The agreement on the low-density tail between the $N$-body realization and the SC and MUSCLE realizations is remarkable. Note even the unphysical low-density peak in 2LPT. At high densities, SC and MUSCLE also perform the most accurately of the approximations, although they still fall well short of $N$-body. The MUSCLE PDF is the closest of the approximations, evidently since it captures additional large-scale collapses compared to SC. But even 'Perfect $\psi$' falls somewhat short. Adding large-scale 2LPT modes to SC and MUSCLE gives almost no change in these PDF's.

Fig. 5 shows 2D histograms of $\psi$, both as predicted using MUSCLE, and as measured at $z = 0$ in the simulation, versus its linear-theory prediction from the initial conditions, $\psi_{\mathrm{Zeld}}$. The re-
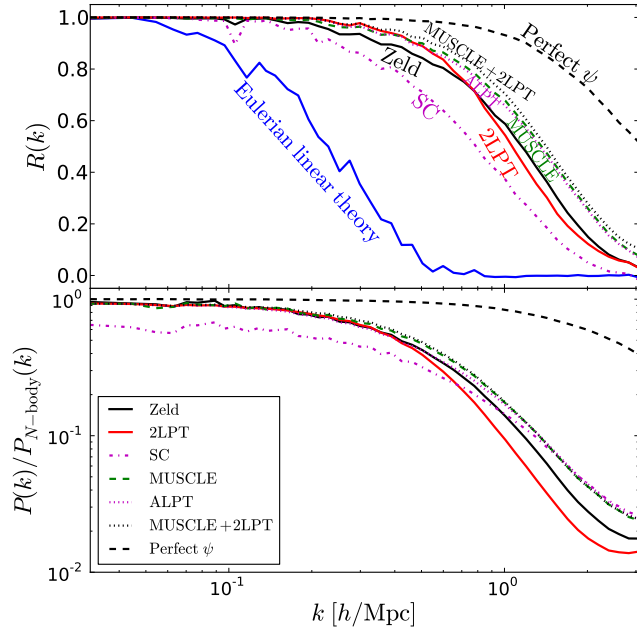
**Figure 3.** *Top:* Fourier-space cross-correlations of the $N$-body density fields with the various approximations. *Bottom:* Ratios of power spectra of the various approximations to that of the full $N$-body density field.
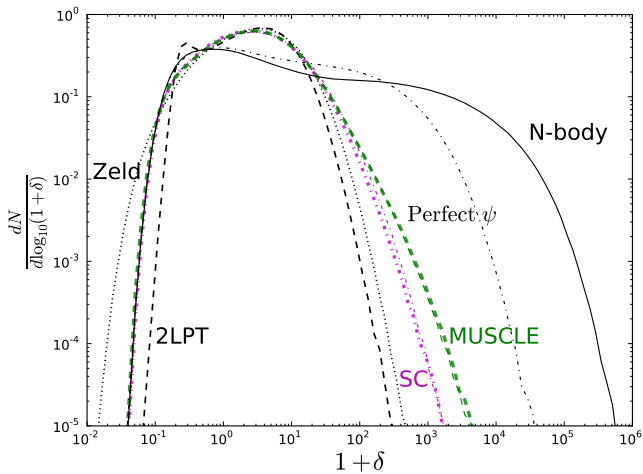


**Figure 4.** Mass-weighted density PDF's of the various realizations, measured using a Voronoi tessellation. The thick magenta dashdotted curve shows the PDF using raw SC, and the thick dashed and curve shows the PDF using MUSCLE. The nearly overlapping thin curves of the same colors show PDF's in ALPT and '2LPT+MUSCLE,' i.e. mixing in 2LPT on large scales.

lationship is simple in MUSCLE: if $\psi_{\rm Zeld} < -1.5$, $\psi_{\rm musc} = -3$. But there are also particles with $\psi_{\rm Zeld} > -1.5$ that get mapped to $-3$; for these, $\psi_{\rm Zeld}$ dips below $-1.5$ when smoothing on some scale. There are two ways used in this paper to measure the divergence $\psi$ from the final conditions: with an FFT; and a 'difference method,' in which $\psi$ is measured at a particle by differencing particle positions initially on either side of it (this has a resolution twice as coarse as the FFT method). The results shown here and in the next figure use the difference method, while the 'Perfect $\psi$' results
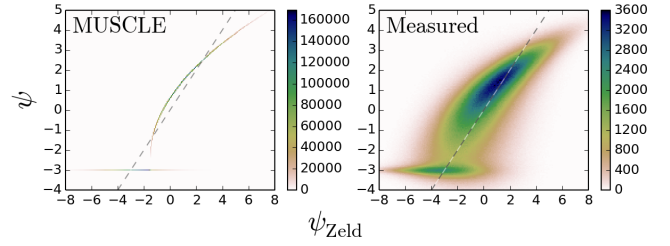


**Figure 5.** A comparison of $\psi = \boldsymbol{\nabla}_q \cdot \boldsymbol{\Psi}$ as predicted from the initial conditions using MUSCLE, to that measured in an $N$-body simulation. $\psi_{\rm Zeld}$, on the $x$-axis of both plots, is the linear-theory (Zel'dovich) prediction. The colorbar numbers indicate the number of particles (out of $256^3$) in that two-dimensional bin. The white-and-black line indicates where the Zel'dovich prediction would lie.

used elsewhere use the FFT method (which does not give a clear locus of points at $\psi = -3$, as seen here).

Another interesting measurement is the distribution of $\psi$ in the $N$-body simulation for different morphologies of particles: void, wall, filament and halo. To classify particles, we use the ORIGAMI algorithm (Falck et al. 2012), which assigns the morphology of a particle according to the number of orthogonal axes by which any other particle crosses it (void=0, wall=1, filament=2, and halo =3), comparing the initial and final conditions.

Fig. 6 shows these PDF's for the various morphologies. Except for 'Halo,' the distributions of $\psi$ look quite Gaussian. One part of the 'Halo' distribution is a sharp spike at $\psi = -3$, characterizing a Lagrangian patch that precisely contracts to an Eulerian point. There is another component that looks somewhat Gaussian, as well; it would be interesting to differentiate these populations physically. One possibility is that the high-$\psi$ tail consists of particles on their first infall. But also, the 'Filament' population has a small bump at $\psi = -3$, suggesting a small amount of contamination between the two morphologies as detected by ORIGAMI.

The positions of the other curves are also informative. From void to halo, the mean $\psi$'s of each morphology are $\psi = (1.80, 0.58, -0.54, -2.25)$. If wall and filament volume elements typically collapsed along one and two axes, their non-collapsing dimensions staying fixed in comoving coordinates, their means would be at $\psi = -1$ and $\psi = -2$. Their higher $\psi$ values indicate that their non-collapsing dimensions tend to stretch substantially.

## 3  DISCUSSION AND CONCLUSION

Here we have presented a conceptually simple, non-perturbative Lagrangian scheme that produces fast $N$-body realizations, essentially as fast as producing initial conditions of a simulation, and more accurately than other such schemes. It applies a spherical-collapse criterion to the Lagrangian divergence of the displacement field, $\psi$, on the pixel scale of the initial conditions, and on larger scales as well. Approximate $N$-body schemes have proven useful to produce mock galaxy catalogs (Kitaura et al. 2014). Some past and upcoming surveys include emission-line galaxies that do not necessarily occupy the largest dark-matter haloes, such as WiggleZ (Drinkwater et al. 2010) and eBOSS (Comparat et al. 2013). For these, fast approximations which accurately resolve the cosmic web to small scales will be increasingly important, for example to produce mock galaxy catalogs with a biasing prescription (e.g. Kitaura et al. 2014; Ata et al. 2015).

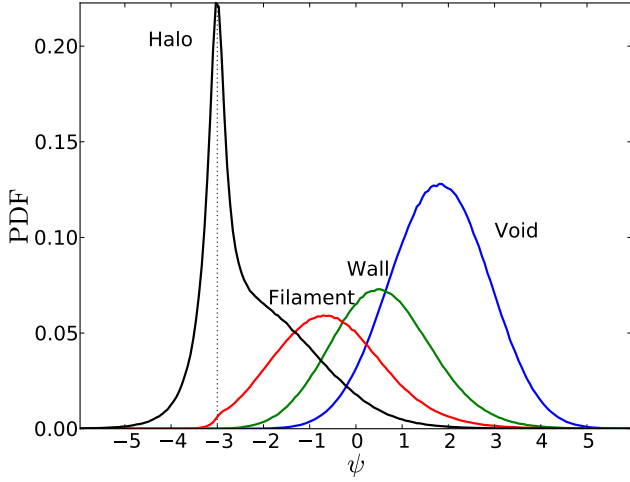There is still room for improvement in such a $\psi$-based scheme;

**Figure 6.** Stretch parameters $\psi = \boldsymbol{\nabla}_q \cdot \boldsymbol{\Psi}$ measured in an $N$-body simulation of interparticle spacing $\sim 1\,h^{-1}$ Mpc, separated by morphological type. The distribution for each type looks rather Gaussian, except for 'Halo;' see discussion in text.

if it were possible to perfectly predict $\psi$, realization accuracies would increase substantially. One possible way forward is to predict the final morphologies (void, wall, filament, halo) of particles, and assign $\psi$ to each differently, according to their values in Fig. 6. Note that although SC and MUSCLE collapse haloes rather well, filaments do not visibly tighten substantially compared to the ZA, at the resolution presented here (at higher resolution, they would tighten some filaments, since MUSCLE limits overcrossing). However, a separate prescription for each morphological type would sacrifice conceptual simplicity, and likely introduce some ad-hoc parameters. Another idea is to remap the approximate realization's matter-density PDF to that of a full $N$-body simulation (Leclercq et al. 2013); note that this step would be bypassed if generating a mock galaxy catalog. Also, we found that interpolating between 2LPT and MUSCLE as in ALPT can slightly improve accuracy over MUSCLE, at the expense of slightly adding to the conceptual complexity and the run time.

A Python package that generates all $\psi$-based particle realizations discussed here, including an ALPT interpolation in scale, is available at http://skysrv.pha.jhu.edu/~neyrinck/muscle.html. It interfaces with CAMB via the COSMOPY package, at http://www.ifa.hawaii.edu/cosmopy/.

## REFERENCES

Abel T., Hahn O., Kaehler R., 2012, MNRAS, 427, 61
Ata M., Kitaura F.-S., Müller V., 2015, MNRAS, 446, 4250
Bernardeau F., 1994, ApJ, 427, 51
Chan K. C., 2014, Phys. Rev. D, 89, 083515
Comparat J., et al., 2013, MNRAS, 428, 1498
Crocce M., Scoccimarro R., 2006, Phys. Rev. D, 73, 063520
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Drinkwater M. J., et al., 2010, MNRAS, 401, 1429
Falck B. L., Neyrinck M. C., Szalay A. S., 2012, ApJ, 754, 126
Heß S., Kitaura F.-S., Gottlöber S., 2013, MNRAS, 435, 2065
Hidding J., van de Weygaert R., Vegter G., Jones B. J. T., 2015, in Rosquist K., ed., Thirteenth Marcel Grossmann Meeting: On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories. pp 2142–2144, doi:10.1142/9789814623995˙0380
Jasche J., Wandelt B. D., 2013, ApJ, 779, 15
Kitaura F.-S., 2013, MNRAS, 429, L84
Kitaura F. S., Enßlin T. A., 2008, MNRAS, 389, 497
Kitaura F.-S., Heß S., 2013, MNRAS, 435, L78
Kitaura F.-S., Erdoğdu P., Nuza S. E., Khalatyan A., Angulo R. E., Hoffman Y., Gottlöber S., 2012, MNRAS, 427, L35
Kitaura F.-S., Yepes G., Prada F., 2014, MNRAS, 439, L21
Kofman L. A., Shandarin S. F., 1988, Nature, 334, 129
Kofman L., Pogosyan D., Shandarin S. F., Melott A. L., 1992, ApJ, 393, 437
Leclercq F., Jasche J., Gil-Marín H., Wandelt B., 2013, J. Cosmology Astropart. Phys., 11, 48
Leclercq F., Jasche J., Wandelt B., 2015, J. Cosmology Astropart. Phys., 6, 15
Mohayaee R., Mathis H., Colombi S., Silk J., 2006, MNRAS, 365, 939
Monaco P., Theuns T., Taffoni G., 2002, MNRAS, 331, 587
Monaco P., Sefusatti E., Borgani S., Crocce M., Fosalba P., Sheth R. K., Theuns T., 2013, MNRAS, 433, 2389
Neyrinck M. C., 2013, MNRAS, 428, 141
Protogeros Z. A. M., Scherrer R. J., 1997, MNRAS, 284, 425
Sahni V., Shandarin S., 1996, MNRAS, 282, 641
Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29
Scoccimarro R., 1998, MNRAS, 299, 1097
Shandarin S. F., 2009, preprint, (arXiv:0912.4520)
Shandarin S., Habib S., Heitmann K., 2012, Phys. Rev. D, 85, 083005
Sheth R. K., van de Weygaert R., 2004, MNRAS, 350, 517
Tassev S., Zaldarriaga M., 2012, J. Cosmology Astropart. Phys., 4, 13
Tassev S., Zaldarriaga M., Eisenstein D. J., 2013, J. Cosmology Astropart. Phys., 6, 36
Valageas P., Bernardeau F., 2011, Phys. Rev. D, 83, 043508
White M., 2014, MNRAS, 439, 3630
White M., Tinker J. L., McBride C. K., 2014, MNRAS, 437, 2594
Zel'dovich Y. B., 1970, A&A, 5, 84